Research Paper | Neuroscience | E-ISSN No : 2454-9916 | Volume : 9 | Issue : 6 | June 2023

PENROSE-HAMEROFF ORCHESTRATED OBJECTIVE REDUCTION THEORY: WHY CONSCIOUSNESS CANNOT BE EFFECTIVELY COMPREHENDED WITH CURRENT LAWS OF PHYSICS

Mos Theodor Alexandru

Research Scholars Program, Harvard Student Agencies, In collaboration with Learn with Leaders

ABSTRACT

This research's purpose is to analyse the Orch OR theory from varying standpoints. Constructive and destructive progress is anatomised within this paper, providing arguments formulated over the years since the hypothesis's inception to offer a synopsis of progress in the field of neurophysics. The methods and conclusions of relevant studies are presented and linked to the final section, wherein the research discusses the necessity for novel theories of physics that describe elaborate events such as those observed within the brain.

KEYWORDS: neurophysics, quantum theory, superposition, consciousness, non-computable systems, microtubules.

INTRODUCTION

Consciousness has long been regarded as labyrinthine and involuted in nature. The path to understanding it is arduous and led to the creation of many a mathematical model describing its existence. Most of these, however, such as the Von Neumann – Wigner interpretation, fell into obscurity as pseudoscientific posits, seemingly veiling every possible lead toward understanding it. Despite this, research in recent years has shown great promise as events occurring in the brain are better understood than ever before. Nevertheless, an issue prevails due to the still novel link between physics and neuroscience as current theories generate dissent within the ranks of researchers.

A particularly controversial postulate is that of Roger Penrose and Stuart Hameroff, which states that the standard model of the brain alone does not account for the presence of consciousness and that awareness is strictly dependent on adjunct quantum-computation elements (McKemmish et al., 2009) attributed to microtubules – cylindrical polymers of "tubulin" protein which organise the internal activity of neurons (Hameroff, 2020).

THEORY DEVELOPMENT

Prior to the nascence of the Orch OR Theory, Penrose (1989) focused on the Czech-born mathematician Kurt Gödel's incompleteness theorem. In 1931, he demonstrated that within any given branch of mathematics, statements would always appear that could not be corroborated or disproven using the rules and axioms of the respective mathematical branch. Implicitly, all logical systems of any complexity are, by definition, incomplete, given that each one of them contains, at any given time, more true statements than it can possibly prove with respect to its own defining set of rules (Denton, 2014). In this regard, Gödel argued that a computer could never be as intelligent as a human being because a fixed set of axioms limits the extent of its knowledge. In contrast, people can discover unexpected truths and expand their collective knowledge ad infinitum.

This theorem acted as the basis of Penrose's first literary work on consciousness, *The Emperor's New Mind* (1989), in which he propounded the idea that Gödel-unprovable results are demonstrable by human mathematicians. In this sense, such individuals cannot be described as formal proof systems and are, therefore, defined by a non-computable algorithm. This affirmation became known as the Lucas-Penrose argument and is comprised of ideas expounded by philosopher J. R. Lucas in his 1961 work *Minds, Machines and Gödel*. Penrose further expanded the boundaries of the claim by viewing the collapse of a wave function – a paramount process in quantum mechanics whereby a particle found in multiple superposed states becomes defined by a classical state upon interaction with its environment – as a non-computable process. This deduction was made after analysing the spontaneity and unpredictability of wave function collapse and was directly linked to his hypothesis regarding non-computable processes found in the brain. As the uncertainty of the event was a tentative physical basis from a mathematical standpoint, Penrose defined another form of collapse for remote systems named objective reduction – OR. The greater the mass-energy of the object, the faster it will undergo OR and vice versa (Penrose R., 2014), a correlation defined by Penrose's indeterminacy principle:

$$\tau \approx \hbar / E_G$$

SOURCE: PENROSE R. AND HAMEROFF S., CONSCIOUSNESS IN THE UNIVERSE (2014)

Concomitantly, Hameroff (1982) studied the properties of anaesthesia – namely, the impact of anaesthetics upon a patient's consciousness – and contributed to cancer research. He was particularly intrigued by the cytoskeleton of neural cells and, by implication, microtubules, which he found were not only providing structural support but also controlling the cell's shape, movement, and growth and played an essential role in axoplasmic transport (Hameroff S., 1987) – that is, the transport of important constituents of cells within the plasmatic body of a neuron (Abe T., Haga T., Kurokawa M., 1973). This led to the conception of Hameroff's germane hypothesis regarding the usage of quantum processing in microtubules, which relies mainly on the nature of tubulin protein dimers. He suggests that the hydrophobic pockets observed in tubulin portions contain delocalised π electrons organised in rings separated by a roughly 2nm gap, which is, in his view, sufficiently close for entanglement to occur (Hameroff S., 2008).

Orch OR coalesces the ideas of Penrose and Hameroff via a link established between the Penrose-Lucas argument and the aforementioned postulate. Penrose and Hameroff suggest that microtubules present superposition states and rely on quantum computation to complete their tasks efficiently (Hameroff S., 1998). Both argue that these structures "orchestrated quantum superpositions, encoding inputs and memory as entangled qubits of collective quantum dipole oscillations". In 1994, Penrose published the Orch OR theory in his book, *Shadows of the Mind*, where he further enforces the non-computable nature of the human brain.

Albeit contentious, the theory was able to maintain relevancy over the years. In 2022, Michael Staelens et al. reported the results of a concatenation of experiments "involving photobiomodulation (PBM) of living cells, tubulin, and microtubules in buffer solutions exposed to near-infrared (NIR) light emitted from an 810 nm LED with a power density of 25 mW/cm^2 pulsed at a frequency of 10 Hz.". Part of these investigations assessed the stability and polymerisation of microtubules under exposure to PBM (Staelens M. et al., 2022). The protein buffer solution used was a mixture of Britton-Robinson buffer (BRB aka PEM) and microtubule cushion buffer. The microtubules began to gradually disassemble after 120 minutes of exposure, an effect later observed in fluorescence microscopy images (Staelens M. et al., 2022). Affiliate researcher Jack Tuszyński also observed that anaesthetics hastens the rate at which trapped light is re-emitted by tubulins and microtubules. He posits that superradiance explains these findings and connects the nature of such structures to quantum effects. Both studies offer a modicum of support to the Orch OR theory.

ISSUES & COUNTERARGUMENTS

Despite the complexity of the theory, most philosophers and scientists deem it overly jejune and explanatorily feeble. The Penrose-Lucas argument received criticism from mathematicians – particularly from Geoffrey LaForte et al. in a study which shows why arguments provided by Penrose erringly rely on "ambiguities between precise and imprecise senses of key terms" (1998) – and philosophers like George Boolos (1990). In his 1991 work *Conscious Machines*, Marvin Minsky argues against Penrose's postulate by stating that the ability of humans to perceive false ideas as veridical demonstrates the inconsistency of human mathematical understanding, thus making consciousness a deterministic process. Similarly, Solomon Feferman argued that the presented comparison between a human and a machine is flawed since the former employs a technique reliant on trial and error rather than automated processes.

Following the publishing of *Shadows of the Mind*, more inconsistencies were found as independent issues or correlated to the Penrose-Lucas argument. The vast majority of counterarguments define the impossibility of such events occurring in biological systems. Physicist Max Tegmark composed the "warm, wet, and noisy" argument, whereby the conditions determined by brain matter would not generate quantum effects as Hameroff believed and would instead nullify them almost instantaneously, not being able to influence neural events (2000). Tegmark concludes that the degrees of freedom of the human brain ought to be seen as classical systems due to the decoherence timescales of microtubule entanglement. Experiments have shown that they are much too short ($\sim 10^{-15}$) seconds) to impact the inner workings of the brain (Tegmark M., 2000) and the respective molecules are much too large to create coherent quantum states (Hossenfelder S., 2021). Christof Koch and Klaus Hepp also concur that a theory of quantum consciousness is not needed for a classical physiological system that has yet to be proven unsound. Koch & Hepp posit that "the empirical demonstration of slowly decoherent and controllable quantum bits in neurons connected by electrical or chemical synapses, or the discovery of an efficient quantum algorithm for computations performed by the brain, would do much to bring these speculations from the 'far-out' to the mere 'very unlikely'" (2006).

Hagan, Tuszynski and Hameroff offered a response to Tegmark's claims, propounding that he did not refer to Orch OR but rather to an altered version of the theory where quanta superpositions were separated by larger gaps than those hypothesised by Penrose and Hameroff. The two argued that microtubule lattice structure enables quantum error correction, which renders decoherence obsolete (Hagan S. et al., 2002). These, however, are logically weak arguments that can be easily countered and enhance the theory's idealistic character.

In 2009, Reimers and McKemmish et al. released a well-defined critique. McKemmish et al. argued that "no reformation of the proposal based on known physical paradigms could lead to quantum computing within microtubules", whilst the other group observed that it is unfeasible for microtubules to support coherence frequencies that are higher than 8Mhz. Hence, it would be impossible for tubulin electrons to form a Bose-Einstein or Frohlich condensate, essential to the events described by Penrose and Hameroff. The foundation of their claims was effectively refuted.

Issues arise in the field of neuroscience as well. Initially, Orch OR relied on 'gap junctions' between glial cells and neurons (Hameroff, 1998). Yet, Binmöller et al. demonstrated in 1992 that this ilk of connection does not exist in the adult brain. In vitro research conducted on rat embryos did, indeed, prove the existence of gap junctions in the early stages of neural development (Fróes et al., 1999). Orch OR, on the other hand, refers to mature cells that do not present this type of bond, therefore contradicting the well-documented properties of neurons and astrocytes. This observation is enforced by Fróes et al. in a 1999 paper: "Junctional communication may provide metabolic and electrotonic interconnections between neuronal and astrocytic networks at early stages of neural development and such interactions are weakened as differentiation progresses." Another relevant biology-related criticism underlines a dearth of explanations regarding the probabilistic release of neurotransmitters from axon terminals.

CONCLUSIONS

Thus, it is not unexpected that the orchestrated Objective Reduction Theory is revered, albeit an ostensibly grand unifying theory between self-awareness and the universe. It regrettably succumbs to its flaws, failing to comprehensively describe the imperfect and hardly definable events that describe consciousness. It is, in essence, a beautiful theoretical model that cannot be applied to real scenarios and that must significantly alter the current paradigm in order to explain its entailments in a given circumstance adequately. One could compare Orch OR to the String Theory or Supersymmetry, both of which are mathematically pleasing but do not account for all the possible implementations of their theses. In spite of this, research which supports the theory is primarily consistent and continually improves its peripheral proprieties. This means it is still possible for it to be proven to some degree, but the consensus among researchers is that such a drastic change in trend is unlikely. Certain individuals believe that consciousness itself cannot be explained in terms of human understanding, and any attempt to 'restrict' it to a physical model is futile.

In this regard, it can be posited that consciousness is essentially incomprehensible at the time due to the paucity of promising theories of physics. To say that something will forever be veiled in obscurity is to ignore inevitable changes in paradigms. Consciousness will eventually be understood, but not with current tools of interpretation. A new theory must be brought forth, one that links biological systems to the general domain of physics. One potential development is reliant on a replacement of quantum mechanics, which would efficiently explain wave function collapse and physical events that occur upon measurement. I suggest it is possible for retrocausality to play a paramount role in determining the nature of consciousness as more and more operations done by the brain are found to be associated with an innate awareness of potential developments – a certain level of prescience, so to speak.

REFERENCES

1. Penrose, R. (1990). The Emperor's New Mind - Concerning Computers, Minds, and the Law of Physics (Book Club (BCE/BOMC)). Oxford.
2. Seth, A. (2021). Being You: A New Science of Consciousness. Dutton.
3. John, E. R. (2002). The neurophysics of consciousness. Brain Research Reviews, 39(1), 1-28. https://doi.org/10.1016/S0165-0173(02)00142-X
4. Sattin, D., Magnani, F. G., Bartesaghi, L., Caputo, M., Fittipaldo, A. V., Cacciatore, M., Picozzi, M., & Leonardi, M. (2021). Theoretical Models of Consciousness: A Scoping Review. Brain Sciences, 11(5). https://doi.org/10.3390/brainsci11050535
5. Koch, C., Massimini, M., Boly, M., & Tononi, G. (2016). Neural correlates of consciousness: progress and problems. Nature Reviews Neuroscience, 17(5), 307–321. https://doi.org/10.1038/nrn.2016.22
6. Stuart, H. (1998). Quantum computation in brain microtubules? The Penrose–Hameroff 'Orch OR' model of consciousness. Philosophical Transactions of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences, 356(1743), 1869–1896. https://doi.org/10.1098/rsta.1998.0254
7. Hameroff, S. (2020). 'Orch OR' is the most complete, and most easily falsifiable theory of consciousness. Cognitive Neuroscience, 12(2), 74–76. https://doi.org/10.1080/17588928.2020.1839037
8. Gödel's Incompleteness Theorem | William Denton. (2014, January 14). https://www.miskatonic.org/godel.html
9. 1931: theoretical computer science & AI theory founded by Goedel. (n.d.). https://people.idsia.ch/%7Ejuergen/goedel-1931-founder-theoretical-computer-science-AI.html
10. Hameroff, S. R., & Watt, R. C. (1982). Information processing in microtubules. Journal of Theoretical Biology, 98(4), 549–561. https://doi.org/10.1016/0022-5193(82)90137-0
11. Abe, T., Haga, T., & Kurokawa, M. (1973). Rapid transport of phosphatidylcholine occurring simultaneously with protein transport in the frog sciatic nerve. Biochemical Journal, 136(3), 731–740. https://doi.org/10.1042/bj1360731
12. Hameroff, S. R. (1987b). Ultimate Computing: Biomolecular Consciousness and Nanotechnology. North-Holland.
13. Staelens, M., Di Gregorio, E., Kalra, A. P., Le, H. T., Hosseinkhah, N., Karimpoor, M., Lim, L., & Tuszyński, J. A. (2022). Near-Infrared Photobiomodulation of Living Cells, Tubulin, and Microtubules In Vitro. Frontiers in Medical Technology, 4. https://doi.org/10.3389/fmedt.2022.871196
14. Tuszynski, J. A., Friesen, D., Freedman, H., Sbitnev, V. I., Kim, H., Santelices, I., Kalra, A. P., Patel, S. D., Shankar, K., & Chua, L. O. (2020). Microtubules as Sub-Cellular Memristors. Scientific Reports, 10(1). https://doi.org/10.1038/s41598-020-58820-y
15. LaForte, G., Hayes, P. J., & Ford, K. M. (1998). Why Gödel's theorem cannot refute computationalism. Artificial Intelligence, 104(1–2), 265–286. https://doi.org/10.1016/s0004-3702(98)00052-6
16. "Conscious Machines", by Marvin Minsky. (n.d.). http://www.aurellem.org/6.868/resources/conscious-machines.html
17. Tegmark, M. (2000). Importance of quantum decoherence in brain processes. Physical Review E, 61(4), 4194–4206. https://doi.org/10.1103/physreve.61.4194
18. Reimers, J. R., McKemmish, L. K., McKenzie, R. H., Mark, A. E., & Hush, N. S. (2009). Weak, strong, and coherent regimes of Fröhlich condensation and their applications to terahertz medicine and quantum consciousness. Proceedings of the National Academy of Sciences of the United States of America, 106(11), 4219-4224. https://doi.org/10.1073/pnas.0806273106
19. Binmöller, F. J., & Müller, C. M. (1992). Postnatal development of dye-coupling among astrocytes in rat visual cortex. Glia, 6(2), 127–137. https://doi.org/10.1002/glia.440060207
20. Hossenfelder, S. (2022, February 19). Has quantum mechanics proved that reality does not exist? [Video]. YouTube. https://www.youtube.com/watch?v=Wsjgtp9XZxo&feature=youtu.be
21. Hossenfelder, S. (2021, January 9). The Mathematics of Consciousness [Video]. YouTube. https://www.youtube.com/watch?v=efVBUDnD_no&feature=youtu.be